Distributed ADMM for in-network reconstruction of sparse signals with innovations

Distributed ADMM for in-network reconstruction of sparse signals with innovations

Javier Matamoros[†], Sophie M. Fosson[⋆], Enrico Magli[⋆] and Carles Antón-Haro[†]
[†]Centre Tecnològic de Telecomunicacions de Catalunya, Barcelona (Spain)
[⋆] Department of Electronics and Telecommunications, Politecnico di Torino (Italy)

***Abstract*—In this paper, we address the problem of in-network reconstruction of correlated sparse signals. Specifically, we adopt a JSM-1 model by which the signals to be reconstructed are the sum of a common sparse term and an individual sparse term (or innovation). We tackle the problem using an Alternating Direction Method of Multipliers approach, which is prone to be distributed. We also propose a version that requires to exchange only binary messages among neighboring nodes. Performance of the different methods is shown to be satisfactory.**

I. Introduction

In recent years, considerable interest has been spurred by compressed sensing theory [1], which states that sparse signals can be reconstructed from a reduced number of observations (measurements). In particular, distributed scenarios have been lately considered [2], [3] where measurements are acquired by a network, and also in-network reconstruction has been addressed to cope with situations where measurements are not available at a central location due to, e.g., privacy reasons or communication constraints [4], [5], [6], [7], [8].

The model for in-network reconstruction typically consists in a number of locally interconnected nodes (or sensors), each of them being associated with a local cost function; the ultimate goal is to minimize the total cost (namely, the sum of the local costs) in a distributed manner. This is accomplished by leveraging on information exchanges over the underlying communication network. To that end, a number of distributed schemes based on the so-called Alternating Direction Method of Multipliers (ADMM, [9]) have been proved to be an attractive solution owing to the recovery accuracy and the fast convergence properties. For example, Mota et al. [4] proposed a distributed ADMM scheme to solve the Basis Pursuit problem, and then generalized it in [10]; parsimonious in terms of communication steps, this algorithm has the drawback of requiring a coloring scheme for the underlying communication graph. In a multi-agent context, Wei and Ozdaglar introduced a distributed ADMM scheme exhibiting a faster convergence rate than traditional subgradient methods [11].

In the aforementioned works, the main goal was to reach consensus on the estimate of a *common* signal. Here, in contrast, we consider a model in which the observed signals to be reconstructed are not identical, but correlated. In particular, we adopt the JSM-1 correlation model [2] whereby each node observation contains a common sparse term plus a sparse innovation (see [3] for possible applications). Centralized reconstruction for JSM-1 has already been addressed in the literature: in [12] some asymptotic bounds were proved, and a single linear program algorithm was used for reconstruction, but complexity was high; in [13], the Texas Hold'Em algorithm was used, which is guaranteed to work when the innovations are incoherent; in [14], [15] side information was exploited for reconstruction. Here, we propose a different approach: we start from the development of a (centralized) ADMM solution, we then propose a distributed version for in-network reconstruction, and finally, in order to reduce the amount of information exchanged that such distributed approach entails, we also propose a novel scheme only requiring binary message exchanges among neighboring nodes. We remark that the choice of considering ADMM for JSM-1 instead of the (centralized) known approaches of [12], [13], [14], [15] is due to ADMM efficiency, mathematical rigor, inclination to be distributed, and ease to be extended to more general multi-agent contexts [11], [16].

This work is supported by European Commission - FP7 Network of Excellence in Wireless COMmunications ♯ (Grant agreement n.318306) and by European Research Council under FP7 / ERC Grant agreement n.279848.

Before proceeding let us introduce some notation. Given $x \in \mathbb{R}^N$, the L^p-norm of x is denoted by $\|x\|_p$ for $p > 0$, whereas $\|x\|_0$ gives the number of non-zero elements of x. The identity matrix of size $L \times L$ will be denoted by I_L. A graph $\mathcal{G}$ is defined as $\mathcal{G} := (\mathcal{N}, \mathcal{E})$ where $\mathcal{N}$ and $\mathcal{E}$ stand for the set of vertices and edges with cardinality $|\mathcal{N}|$ and $|\mathcal{E}|$ respectively.

II. Signal Model

Consider a network composed of N nodes whose connectivity is described through the connected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$. Accordingly, node $i \in \mathcal{N}$ can communicate with node $j \in \mathcal{N}$ if the edge $\{i, j\}$ is included in $\mathcal{E}$, or, in other words, j belongs to the neighborhood set of i, denoted as $\mathcal{N}_i$. In this scenario, each node observes a compressed version of a signal $\{x_i\}_{i \in \mathcal{N}} \in \mathbb{R}^n$ through a set of linear and local measurements, namely

$$y_i = A_i x_i + \eta_i \quad ; \quad i \in \mathcal{N} \tag{1}$$

where $A_i \in \mathbb{R}^{M \times L}$ (with $M \ll L$) stands for the measurement matrix at the *i-th* node and $\eta_i \in \mathbb{R}^L$ for additive noise. We further assume that the observed signals follow the JSM-1 model [3], namely

$$x_i = z_c + z_i \quad ; \quad i \in \mathcal{N}. \tag{2}$$

Algorithm 1 Computation of $z_c(t+1), \{z_i(t+1)\}$

```
1: Initialize 0 < ε ≪ 1 and 0 < δ ≪ 1; initialize auxiliary
   variables z_{c,l}^{(+)} = 0, z_{c,l}^{(-)} = 0, for l = 1, ..., L
2: for l = 1, ..., L do
3:    Assume z_{c,l} = 0; compute z_{i,l}, s_{c,l} by (13),(16)
4:    if s_{c,l} ∈ (−1, 1) then
5:       z_{c,l} ← 0 and keep z_{i,l} from step 3; stop
6:    end if
7:    loop
8:       z_{c,l}^{(+)} ← z_{c,l}^{(+)} + ε and z_{c,l}^{(-)} ← z_{c,l}^{(-)} − ε
9:       Assume z_{c,l} = z_{c,l}^{(+)}; compute z_{i,l}, s_{c,l} by (13), (16)
10:      if ||s_{c,l} − 1||_2^2 < δ then
11:         z_{c,l} ← z_{c,l}^{(+)} and keep z_{i,l} from step 9; stop
12:      end if
13:      Assume z_{c,l} = z_{c,l}^{(-)}; compute z_{i,l}, s_{c,l} by (13), (16)
14:      if ||s_{c,l} + 1||_2^2 < δ then
15:         z_{c,l} ← z_{c,l}^{(-)} and keep z_{i,l} from step 13; stop
16:      end if
17:   end loop
18: end for
```

That is, the observed signal at each node is composed of a common component plus an innovation component. In addition, we consider that z_c and $\{z_i\}$ are both unknown and sparse, with the number of non-zero elements given by $k_c = \|z_c\|_0$ and $\{\|z_i\|_0\} = k$, respectively. As for the signal supports, defined as $\Omega_i := \{l | z_{i,l} \neq 0\}$ for $i \in \mathcal{N}$ and $\Omega_c := \{l | z_{c,l} \neq 0\}$, do not necessarily coincide.

The ultimate goal is to reconstruct the triplets $\{x_i, z_c, z_i\}$ at each node in a distributed manner. To that end, we attempt to solve the following convex optimization problem:

$$\min_{\{x_i, z_i\}, z_c} \frac{1}{2} \sum_{i=1}^{N} \left(\|y_i - A_i x_i\|_2^2 + \tau_1 \|z_i\|_1 + \tau_2 \|z_c\|_1 \right) \quad (3)$$

$$\text{s.t.} \quad x_i = z_c + z_i; \quad i = 1, \ldots, N, \quad (4)$$

with τ_1 and τ_2 denoting weights aimed to promote sparsity in the individual and common components, respectively.

III. ADMM for JSM-1

In this section, we address the centralized reconstruction. In particular, we propose a (centralized) ADMM solution [9], which has not been yet addressed in the literature for JSM-1. We will use it as basis to develop our distributed schemes, which are our main purpose, and as benchmark to test them. Following the ADMM rationale, we augment the cost function in (3) as

$$\min_{\{x_i, z_i\}, z_c} \frac{1}{2} \sum_{i=1}^{N} \Big(\|y_i - A_i x_i\|_2^2 + \tau_1 \|z_i\|_1 + \tau_2 \|z_c\|_1 + \frac{\rho}{2} \|x_i - z_i - z_c\|_2^2 \Big) \quad (5)$$

$$\text{s.t.} \quad x_i = z_i + z_c; \quad i = 1, \ldots, N \quad (6)$$

where ρ is a positive constant. Thus, the Lagrangian of the augmented problem reads:

$$\mathcal{L} := \frac{1}{2} \sum_{i=1}^{N} \|y_i - A_i x_i\|_2^2 + \sum_{i=1}^{N} \tau_1 \|z_i\|_1 + N\tau_2 \|z_c\|_1 + \sum_{i=1}^{N} \frac{\rho}{2} \|x_i - z_i - z_c\|_2^2 + \sum_{i=1}^{N} \lambda_i^T (x_i - z_i - z_c), \quad (7)$$

with $\{\lambda_i\}$ standing for the Lagrangian multipliers associated to the constraints in (18). Hence, the ADMM iterates in the primal and dual domain [9] as follows:

$$x_i(t+1) = (\rho I + A_i^T A_i)^{-1} (A_i^T y_i + \rho(z_i(t) + z_c(t)) - \lambda_i^T(t))$$
$$z_c(t+1), \{z_i(t+1)\} = \arg \min_{z_c, \{z_i\}} \mathcal{L}(t+1) \quad (8)$$
$$\lambda_i(t+1) = \lambda_i(t) + \rho \left(x_i(t+1) - z_i(t+1) - z_c(t+1)\right)$$

where $\mathcal{L}(t+1)$ stands for the Lagrangian of (7) evaluated at $\{x_i(t+1), \lambda_i(t)\}$. As for the minimization step of (8), the solution must satisfy the following system of equations:

$$\partial_{z_i} \mathcal{L}(t+1) = 0; \quad i \in N, \quad (9)$$
$$\partial_{z_c} \mathcal{L}(t+1) = 0, \quad (10)$$

with $\partial_x f$ denoting the subgradient of f with respect to x (see definition in [17]). From (9), the optimal variables $\{z_i(t+1), z_c(t+1)\}$ must satisfy:

$$\tau_1 s_i - \rho(x_i(t+1) - z_i(t+1) - z_c(t+1)) - \lambda_i(t) = 0 \quad (11)$$

for $i \in N$. In the equation above, the L-length vector s_i stands for the subgradient of $\|z_i\|_1$ evaluated at $z_i(t+1)$ and its components are $s_{i,l} = 1$ if $z_{i,l}(t+1) > 0$, $s_{i,l} = -1$ for $z_{i,l}(t+1) < 0$ and

$$s_{i,l} = \frac{\rho}{\tau_1}(x_i(t+1) - z_c(t+1)) - \frac{\lambda_i(t)}{\tau_1} \in (-1, 1) \quad (12)$$

for $z_{i,l}(t+1) = 0$. From all the above and assuming that $z_c(t+1)$ is known, $z_i(t+1)$ reads

$$z_i(t+1) = \mathcal{S}_{\frac{\tau_1}{\rho}} \left[x_i(t+1) - z_c(t+1) + \frac{\lambda_i(t)}{\rho} \right], \quad (13)$$

with $\mathcal{S}_\alpha(a)$ standing for the well-known soft-thresholding operator. That is, $\mathcal{S}_\alpha(a) = a - \alpha$ if $a > \alpha$, $\mathcal{S}_\alpha(a) = a + \alpha$ if $a < -\alpha$ and $\mathcal{S}_\alpha(a) = 0$ otherwise. As for the common component $z_c(t+1)$, let s_c be the subgradient of $\|z_c\|_1$ evaluated at $z_c(t+1)$ with entries given by $s_{c,l} = 1$ if $z_{c,l}(t+1) > 0$, $s_{c,l} = -1$ for $z_{c,l}(t+1) < 0$ and $s_{c,l} \in (-1, 1)$ for $z_{c,l}(t+1) = 0$ for $l = 1, \ldots, L$. Accordingly, we have that

$$\partial_{z_{c,l}} \mathcal{L}(t+1) = N\tau_2 s_{c,l} - \sum_{i=1}^{N} \lambda_i(t) - \sum_{i=1}^{N} \rho(x_i(t+1) - z_i(t+1) - z_c(t+1)) \quad (14)$$

$$= N\tau_2 s_{c,l} - \tau_1 \sum_{i=1}^{N} s_{i,l} \quad ; l = 1, \ldots, L, \quad (15)$$

where the last step follows from (11). Finally, from (10), the subgradient of $\|z_{c,l}(t+1)\|_1$ must satisfy:

$$s_{c,l} = \frac{\tau_1}{N\tau_2}\sum_{i=1}^{N} s_{i,l} \in [-1,1] \quad ; \quad l = 1,\ldots,N. \tag{16}$$

Bearing all the above in mind, we propose Algorithm 1 to find the set of $z_c(t+1), \{z_i(t+1)\}$ that solve (8). First, we assume that $z_{c,l} = 0$, $l = 1,\ldots,L$, then we obtain all individual components $z_{i,l}$ from (13), and check whether (16) holds true. If so, we conclude that $z_{c,l} = 0$ and terminate. Otherwise, the algorithm performs a line search over $z_{c,l}$ by progressively increasing/decreasing the positive and negative guesses on $z_{c,l}$ until the $s_{c,l}$ computed in Steps 9 and 13 matches the current guesses ($+1$ or -1, for $z^+_{c,l}$ or $z^+_{c,l}$).

IV. Distributed ADMM for JSM-1

This section goes one step beyond Section III and attempts to find a *distributed* reconstruction method. To that end, we propose to solve the following optimization problem:

$$\min_{\{x_i,z_i,\zeta_i,c_i\}} \frac{1}{2}\sum_{i=1}^{N}\|y_i - A_i x_i\|_2^2 + \tau_1\|z_i\|_1 + \tau_2\|\zeta_i\|_1 \tag{17}$$

$$\text{s.t.} \quad x_i = z_i + \zeta_i; \quad i \in N \tag{18}$$

$$\zeta_i = c_j; \quad j \in \mathcal{N}_i \tag{19}$$

Here, we have introduced the local variables $\{\zeta_i\}$, $\{c_i\}$ that must be interpreted as the local and neighbors guesses on the common component. The consensus constraint of (19) and the fact that $\mathcal{G}$ is a connected graph make the problem above still equivalent to (3). In order to solve (17)–(19), we resort again to the ADMM and build the following augmented cost function:

$$\min_{\{x_i,z_i,\zeta_i,c_i\}} \frac{1}{2}\sum_{i=1}^{N}\|y_i - A_i x_i\|_2^2 + \tau_1\|z_i\|_1 + \tau_2\|\zeta_i\|_1 + \frac{\rho}{2}\|x_i - \zeta_i\|_2^2 + \frac{\theta}{2}\sum_{j\in\mathcal{N}_i}\|\zeta_i - c_j\|_2^2 \tag{20}$$

$$\text{s.t.} \quad x_i = z_i + \zeta_i; \quad i \in N \tag{21}$$

$$\zeta_i = c_j; \quad j \in \mathcal{N}_i \tag{22}$$

with ρ and θ standing for positive constants. Now, in an attempt to find a distributed the solution for (20), we propose to sequentially update the primal variables $\{x_i, z_i, \zeta_i, c_i\}$ according to

$$x_i(t+1) = (\rho I + A_i^T A_i)^{-1}(A_i^T y_i + \rho(z_i(t) + \zeta_i(t)) - \lambda_i^T)$$

$$z_i(t+1) = \mathcal{S}_{\frac{\tau_1}{\rho}}\left[(x_i(t+1) - \zeta_i(t)) + \frac{\lambda_i(t)}{\rho}\right]$$

$$\zeta_i(t+1) = \mathcal{S}_{\frac{\tau_2}{\rho+\theta|\mathcal{N}_i|}}\left[\frac{1}{\rho+\theta|\mathcal{N}_i|}\left(\rho\left(x_i(t+1) - z_i(t+1)\right) + \theta\sum_{j\in N_i}\left(c_j(t) - \frac{\mu_{i,j}(t)}{\theta}\right) + \lambda_i(t)\right)\right] \tag{23}$$

$$c_i(t+1) = \frac{1}{|\mathcal{N}_i|}\sum_{j:i\in\mathcal{N}_j}\left(\zeta_j(t+1) + \frac{\mu_{j,i}(t)}{\theta}\right); \tag{24}$$

followed by the ascent updates of the dual variables, that is,

$$\lambda_i(t+1) = \lambda_i(t) + \rho\left(x_i(t+1) - z_i(t+1) - \zeta(t+1)\right)$$

$$\mu_{i,j}(t+1) = \mu_{i,j}(t) + \theta\left(\zeta_i(t+1) - c_j(t+1)\right); \quad j \in \mathcal{N}_i,$$

where $\{\lambda_i\}$ and $\{\mu_{i,j}\}$ stand for the Lagrangian multipliers associated to constraints (21) and (22) respectively. Interestingly, this iterative method can be readily implemented in a distributed manner by exchanging information among neighbor nodes only. The proposed distributed ADMM for JSM-1 (referred to in the sequel as DADMM), is summarized in Algorithm 2. Notice that DADMM retrieves rationale of the algorithm proposed in [11], but extends to the case of different (though correlated) signals. On the other hand, we believe that the proof of convergence in [11] could be exploited to prove the convergence of DADMM, which will be the main focus of a future extended work.

Algorithm 2 Distributed ADMM (DADMM)

for all $i \in \mathcal{N}$ **do**
    Initialize variables: $x_i(1) = [0,0,\ldots,0]^T$; $z_i(1) = [0,0,\ldots,0]^T$; $\zeta_i(1) = [0,0,\ldots,0]^T$; and $c_i(1) = [0,0,\ldots,0]^T$
end for
for $t = 1,\ldots,T_{\max}$ **do**
    for all $i \in \mathcal{N}$ **do**
        $x_i(t+1) \leftarrow (\rho I + A_i^T A_i)^{-1}(A_i^T y_i + \rho(z_i(t)+\zeta_i(t)) - \lambda_i(t)^T)$
        $z_i(t+1) \leftarrow \mathcal{S}_{\frac{\tau_1}{\rho}}\left[(x_i(t+1) - \zeta_i(t)) + \frac{\lambda_i(t)}{\rho}\right]$
        $\zeta_i(t+1) \leftarrow \mathcal{S}_{\frac{\tau_2}{\rho+\theta|\mathcal{N}_i|}}\left[\frac{1}{\rho+\theta|\mathcal{N}_i|}\left(\rho\left(x_i(t+1) - z_i(t+1)\right) + \theta\sum_{j\in N_i}\left(c_j(t) - \frac{\mu_{i,j}(t)}{\theta}\right) + \lambda_i(t)\right)\right]$
        Broadcast $\zeta_i(t+1)$ to each node j with $j \in \mathcal{N}_i$
        $c_i(t+1) \leftarrow \frac{1}{|\mathcal{N}_i|}\sum_{j:i\in\mathcal{N}_j}\left(\zeta_j(t+1) + \frac{\mu_{j,i}(t)}{\theta}\right)$
        Broadcast $c_i(t+1)$ to each node j with $j \in \mathcal{N}_i$
        $\lambda_i(t+1) \leftarrow \lambda_i(t) + \rho\big(x_i(t+1) - z_i(t+1) - \zeta(t+1)\big)$
        for all $j \in \mathcal{N}_i$ **do**
            $\mu_{i,j}(t+1) \leftarrow \mu_{i,j}(t+1) + \theta\left(\zeta_i(t+1) - c_j(t+1)\right)$
        end for
        for all $j : i \in \mathcal{N}_j$ **do**
            $\mu_{j,i}(t+1) \leftarrow \mu_{j,i}(t) + \theta\left(\zeta_j(t+1) - c_i(t+1)\right)$
        end for
    end for
end for

V. Distributed ADMM with 1 bit messages

The main drawback of the proposed DADMM scheme is the large amount of information that needs to be exchanged among neighboring nodes (i.e. $\zeta_i(t+1)$ and $c_i(t+1)$, in each iteration). This in turn results in a large energy consumption and reduced network lifetime. To circumvent that, we propose to quantize the exchanged variables with 1 bit only. In order to retain most of the advantages of the scheme, we replace the primal updates of (23) and (24) (steps 8 and 10 in Algorithm 2) by gradient updates of constant step length ϵ, that is

$$\zeta_i(t+1) = \zeta_i(t) - \epsilon \frac{g_{\zeta_i^t}}{\|g_{\zeta_i^t}\|_1} = \zeta_i(t) - \epsilon \,\mathrm{sign}\left(g_{\zeta_i^t}\right) \quad (25)$$

$$c_i(t+1) = c_i(t) - \epsilon \frac{g_{c_i^t}}{\|g_{c_i^t}\|_1} = c_i(t) - \epsilon \,\mathrm{sign}\left(g_{c_i^t}\right), \quad (26)$$

where $g_{\zeta_i^t}$ and $g_{c_i^t}$ stand for the subgradient of the augmented Lagrangian with respect to ζ_i and c_i at time t and; $\mathrm{sign}(x) = 1$ if $x \geq 0$ and $\mathrm{sign}(x) = -1$ otherwise. Consequently, in DADMM-1bit nodes only need to broadcast the sign of the innovations, namely $\mathrm{sign}(g_{\zeta_i^t})$ and $\mathrm{sign}(g_{c_i^t})$, in steps 9 and 11 of Algorithm 2.

As for the computation of $g_{\zeta_i^t}$, note that

$$g_{\zeta_i^t} := \tau_2 s - \rho(x_i(t+1) - z_i(t+1) - \zeta_i(t)) - \lambda_i(t) + \sum_{j \in \mathcal{N}_i} \theta\left(\zeta_i(t) - c_j(t)\right) + \mu_{i,j}(t), \quad (27)$$

where the L-length vector s stands for the subgradient of $\|\zeta_i(t)\|_1$. Similarly, for $g_{c_i^t}$ we have that

$$g_{c_i^t} := -\theta \sum_{j: i \in \mathcal{N}_j} \left(\zeta_j - c_i(t) - \frac{\mu_{j,i}}{\theta}\right). \quad (28)$$

VI. Numerical Results and Conclusions

In the simulations, we consider noiseless measurements and signals $\{z_c, z_i\}$ of length $L = 100$ with sparsity levels $\{k_i\} = k_c = 5$. The supports of the individual and common signals $\{z_c, z_i\}$ are generated uniformly at random, with non-zero elements drawn from a standard Gaussian distribution. As a performance metric, we use the normalized mean square error, which for a generic k-sparse signal x is defined as $\mathrm{MSE}(x) = \frac{1}{k\sigma_x^2}\|\hat{x} - x\|_2^2$ with $\hat{x}$ standing for the reconstruction of x and σ_x^2 for the average power of the non-zero values.

In Figure 1, we plot the attained MSE for the three proposed reconstruction methods (DADMM-1bit is tested for different values of ϵ defined in (25)). In this setting, we have considered $M = 25$ and a regular graph with degree $d = 5$ for the distributed cases. Unsurprisingly, the centralized approach converges much faster than its distributed counterparts. Still, both the DADMM and the DADMM-1bit with $\epsilon = 0.01$ also achieve perfect reconstruction. For DADMM-1bit, we observe that ϵ impacts on the accuracy of the estimates and on the convergence speed: when ϵ increases, the algorithm converges faster at the price of less accurate estimation. Besides, this also explains the MSE oscillations for large values of ϵ. More interestingly, for small values, like $\epsilon = 0.01$, DADMM-1bit performs virtually identical to DADMM at the expense of 3 times more iterations to converge. From a signalling viewpoint this is still favorable: if, for instance, real values can be quantized over 16 bits, the signalling ratio is 3/16.

Finally, Figure 2 shows the attained MSE in the reconstruction of the individual (z_i) and common signals (z_c). In these simulations, we have considered a case with a lower number of measurements per node (i.e. $M = 20$) which does not allow perfect reconstruction of $\{x_i\}$. Interestingly, all algorithms achieve perfect reconstruction of the common component z_c thanks to the redundancy in the number of node measurements but are unable to reconstruct the innovations. Again, all exhibit an identical performance after convergence.

In summary, we have addressed the problem of in-network reconstruction of sparse signals with innovations. As a result, we have proposed two distributed ADMM schemes, that are shown to converge to the centralized ADMM solution in a reasonable number of iterations. The 1-bit version is shown to reduce the total number of transmitted bits. Future work will envisage comparisons to known centralized solutions for JSM-1, convergence analysis and extension to other correlation models.

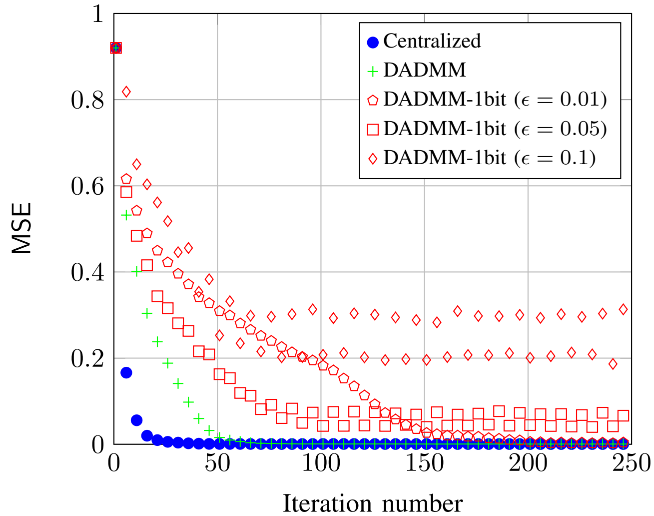

Figure 1. MSE in the reconstruction of $\{x_i\}$ vs iteration number ($N = 20$, $M = 25$, $k_i = k_c = 5$, $L = 100$, $d = 5$, $\tau_1 = 3 \cdot 10^{-3}$, $\tau_2 = 6 \cdot 10^{-4}$, $\rho = \theta = 0.01$).

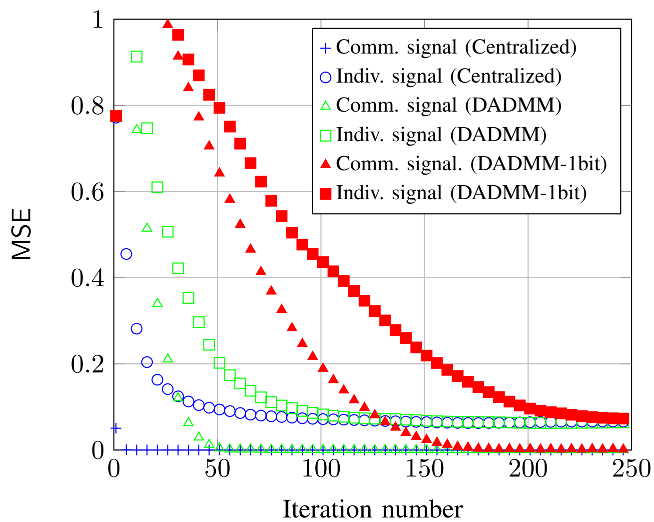

Figure 2. MSE in the reconstruction of $\{z_i\}$ and z_c vs iteration number ($N = 20$, $M = 20$, $k_i = k_c = 5$, $L = 100$, $d = 5$, $\tau_1 = 3 \cdot 10^{-3}$, $\tau_2 = 6 \cdot 10^{-4}$, $\rho = \theta = 0.01$).

References

[1] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] M. F. Duarte, S. Sarvotham, M. B. Wakin, D. Baron, and R. G. Baraniuk, "Joint sparsity models for distributed compressed sensing," in *Online Proceedings of the Workshop on Signal Prcessing with Adaptive Sparse Structured Representations (SPARS)*, 2005.

[3] M. F. Duarte, S. Sarvotham, D. Baron, M. B. Wakin, and R. G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *Proc. 39th Asilomar Conf. Signals, Systems and Computers, Pacific Grove, CA*, 2005.

[4] J. F. C. Mota, J. M. F. Xavier, P. M. Q. Aguiar, and M. Püschel, "Distributed basis pursuit," *IEEE Trans. on Signal Processing*, vol. 60, no. 4, pp. 1942–1956, 2012.

[5] ——, "D-ADMM: A distributed algorithm for compressed sensing and other separable optimization problems," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2012, pp. 2869–2872.

[6] D. Sundman, S. Chatterjee, and M. Skoglund, "Distributed greedy pursuit algorithms," *submitted to EURASIP Signal Processing*, 2013.

[7] C. Ravazzi, S. M. Fosson, and E. Magli, "Distributed soft thresholding for sparse signal recovery," in *Globecom*, 2013.

[8] S. Fosson, J. Matamoros, C. Anton-Haro, and E. Magli, "Distributed support detection of jointly sparse signals," in *Proc. of 39th ICASSP*, 2014.

[9] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2010.

[10] J. F. C. Mota, J. M. F. Xavier, P. M. Q. Aguiar, and M. Püschel, "D-admm: A communication-efficient distributed algorithm for separable optimization," *IEEE Trans. on Signal Processing*, vol. 61, no. 10, pp. 2718–2723, 2013.

[11] E. Wei and A. Ozdaglar, "Distributed alternating direction method of multipliers," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, Dec 2012, pp. 5445–5450.

[12] D. Baron, M. F. Duarte, S. Sarvotham, M. B. Wakin, and R. G. Baraniuk, "An information theoretic approach to distributed compressed sensing," in *Allerton Conference on Communication, Control, and Computing*, Allerton, IL, Sep. 2005.

[13] S. R. Schnelle, J. N. Laska, C. Hegde, M. F. Duarte, M. A. Davenport, and R. G. Baraniuk, "Texas Hold 'Em algorithms for distributed compressive sensing," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, TX, Mar. 2010.

[14] G. Coluccia, E. Magli, A. Roumy, and V. Toto-Zarasoa, "Lossy compression of distributed sparse sources: a practical scheme," in *2011 European Signal Processing Conference (EUSIPCO-2011)*, 2011, pp. 422–426.

[15] D. Valsesia, G. Coluccia, and E. Magli, "Joint recovery algorithms using difference of innovations for distributed compressed sensing," in *Signals, Systems and Computers, 2013 Asilomar Conference on*, Nov 2013, pp. 414–417.

[16] E. Wei and A. Ozdaglar, "On the $o(1/k)$ convergence of asynchronous distributed alternating direction method of multipliers," in *Global Conference on Signal and Information Processing (GlobalSIP), 2013 IEEE*, Dec 2013, pp. 551–554.

[17] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 1993.